

## Paralelização do processo de produção de mapas da RCFM

José Oscar Fernandes<sup>\*,1</sup>, Airam Jônatas Preto<sup>\*\*,1</sup>, Stephan Stephany<sup>\*\*,2</sup>, Carlos Alexandre Wuensche<sup>3</sup>

(1)Área de Computação Científica e Processamento de Alto Desempenho  
Setor de Lançamento de Balões

Instituto Nacional de Pesquisas Espaciais (INPE)

(2)Laboratório Associado de Computação e Matemática Aplicada  
Instituto Nacional de Pesquisas Espaciais (INPE)

(3)Divisão de Astrofísica

Instituto Nacional de Pesquisas Espaciais (INPE)

(\*)Mestrado, e-mail: [oscar@das.inpe.br](mailto:oscar@das.inpe.br); (\*\*) Orientadores

### Resumo

Este trabalho descreve a paralelização de um código para produção de mapas da RCFM (Radiação Cósmica de Fundo em Microondas) em uma arquitetura de memória distribuída. Inicialmente, o código escrito em Fortran 77 foi portado para Fortran 90 e HPF (High Performance Fortran). O código foi instrumentado para análise de tempo de processamento e foi identificado como trecho crítico o algoritmo de convolução de matrizes utilizando 2D-FFT (two-dimensional Fast Fourier Transform). Com a distribuição das matrizes em blocos de colunas entre processadores, a comunicação gerada pela transposição de matrizes é crítica nesta implementação. É proposto o uso de funções MPI (Message Passing Interface) em conjunto com HPF, visando uma comunicação mais eficiente entre processadores.

*Palavras-Chave: Fortran 90, HPF, MPI, paralelismo de dados, análise de dados, Radiação Cósmica de Fundo*

### Introdução

A RCFM é uma prova a favor do modelo da grande explosão, ou Big Bang, como evento responsável pelo surgimento do universo. A energia liberada após a explosão teria se arrefecido com a expansão do Universo, fazendo com que a temperatura do Universo tenha também decrescido proporcionalmente à expansão. A observação desta energia residual é uma ferramenta importante para estudo dos mecanismos responsáveis pela formação de estruturas (galáxias, aglomerados) do Universo.

Telescópios embarcados em balões estratosféricos e satélites têm sido utilizados para estudo da RCFM. Medidas de temperatura são feitas repetidamente em uma determinada região do céu, gerando uma série temporal de dados, a partir dos quais são gerados, sucessivamente, mapas do céu, estimativas do espectro de potência das flutuações de temperatura e estimativa de parâmetros cosmológicos.

Para produção de mapas, um conjunto de sub-vetores desta série é convenientemente definido e o método dos gradientes conjugados é aplicado a cada sub-vetor para compor a matriz de covariância de ruído. A aplicação de um filtro "white-noise" a esta matriz produz o mapa da RCFM.

O grande volume de dados, atingindo de  $10^6$  a  $10^7$  pontos, e a incapacidade dos algoritmos atuais em lidar com as técnicas de análise em escalas de tempo adequadas motivaram o estudo de implementações alternativas e a paralelização destes procedimentos.

Este trabalho envolve a reestruturação e otimização do código utilizando Fortran 90 e a paralelização do processo utilizando HPF, tendo como plataforma um "cluster" de 16 microcomputadores Pentium III de 800MHz, rodando Linux (kernel 2.4.2), com 4 GB de memória distribuída e 320 GB de área para armazenamento de dados. Em trechos do programa onde o modelo baseado em paralelismo de dados se mostra ineficiente, como no cálculo de matrizes transpostas, é proposta a utilização de funções MPI (Message Passing Interface), esperando-se minimizar a comunicação entre processadores.

### Processo de Produção de Mapas da RCF

A implementação atual foi desenvolvida para o experimento HACME (HEMT Advanced Cosmic Microwave Explorer), cujo objetivo foi medir flutuações de temperatura da RCF em regiões fora do plano da Galáxia. Paralelamente o HACME serviu de protótipo para testar a instrumentação que fará parte do experimento ACE (Advanced Cosmic Explorer). Este experimento consiste em um telescópio gregoriano

não-axial formado por espelho primário paraboloïdal (semi-eixo maior = 2,2 m) e um sub-refletor elipsoidal (semi-eixo maior = 0,8 m). Um espelho plano giratório faz com que o feixe descreva uma trajetória elíptica no céu que, combinada com o movimento azimutal da gôndola, possibilita a obtenção de mapas bidimensionais da RCFM. O telescópio é montado em uma plataforma controlada e fará vôos de cerca de 100 dias de duração a bordo de balões estratosféricos superpressurizados. Ao final da missão, com 4 vôos em regiões distintas, o experimento terá mapeado cerca de 75% da esfera celeste com uma sensibilidade jamais atingida por outro experimento.

Chama-se mapa da RCFM a representação da distribuição da intensidade de brilho do céu obtida através dos valores da temperatura de cada um dos pixels da região observada. Conhecendo-se a posição em que o telescópio estava apontado em cada observação  $(\theta_t, \psi_t)$  e adotando uma divisão do céu em pixels, pode-se construir uma matriz de apontamento  $P$  dada pela intensidade do pixel  $p$  na observação  $t$ , sendo  $P=1$  se  $(\theta_t, \psi_t) \in p$  e  $P=0$ , caso contrário.

O vetor de dados pode ser escrito como  $d=Ps+n$ , onde  $s$  é o sinal desconhecido do céu,  $n$  o ruído (normalmente "white noise") e  $P$  a matriz de apontamento.

O mapa é representado pela equação:  $m=(P^t N^t P)^{-1} P^t N^t d$ , onde  $N=\langle nn^t \rangle$  é a matriz de correlação de ruído tempo-tempo.

Substituindo o vetor  $d$  por sua equação, tem-se que:  $m=s+(P^t N^t P)^{-1} P^t N^t n$ .

O ruído de cada pixel é dado por:  $v=(P^t N^t P)^{-1} P^t N^t n$ , com correlações dadas por:  $\Upsilon=\langle vv^t \rangle=(P^t N^t P)^{-1}$ .

O processo de produção de mapas é convenientemente dividido em 3 etapas, que envolvem a solução das equações:  $\Upsilon^{-1}=P^t N^t P$ ,  $z=P^t N^t d$  e  $m=(\Upsilon^{-1})^{-1} z$ .

Como o vetor que contém a série temporal é geralmente muito maior que a memória disponível, o processo é feito em seções. Cada seção é lida e convoluída com o filtro "white noise".

A série temporal é multiplicada por  $P^t$  e convertida para pixels para cada seção, sendo o processo repetido em toda a extensão do vetor, usando o método "overlap-add" para a convolução.

A Tabela 1 apresenta estimativa de recursos computacionais requeridos para cada passo do processo explorando-se as estruturas dos dados.  $N_p$  é o número de pixels do mapa e  $N_t$  o número de pontos da série temporal. A matriz de apontamento é esparsa com o número de entradas não-zero em cada linha da matriz de apontamento ( $N_a$ ) tipicamente igual a 0, 1 ou 2. As matrizes são simétricas e o ruído de correlação temporal estacionário tende a zero, devido ao tempo de separação ser muito menor que o de observação, sendo  $N_{tt}^{-1} = f(|t-t'|) = 0 \forall |t-t'| > \tau \ll N_t$ .

|                            | Disco    | RAM      | Operações        |
|----------------------------|----------|----------|------------------|
| $\Upsilon^{-1}=P^t N^t P$  | $4N_p^2$ | $8N_p^2$ | $3N_a^2 N_t N_t$ |
| $z=P^t N^t d$              | $4N_t$   | $8N_t$   | $3N_a N_t N_t$   |
| $m=(\Upsilon^{-1})^{-1} z$ | $4N_p^2$ | $8N_p^2$ | $(2+2/3)N_p^3$   |

Tabela 1 - Estimativa de recursos computacionais

### Resultados parciais

O código referente à primeira etapa do processo de produção de mapas, ou seja, o cálculo da matriz de covariância de ruído, foi otimizado obtendo-se um ganho de 35% no tempo de processamento. Foram identificados como trechos críticos do código as sub-rotinas de cálculo de convolução de matrizes utilizando 2D-FFT, que são realizadas através dos seguintes passos:

- Cálculo da 2D\_FFT da matriz A
- Cálculo da 2D\_FFT da matriz B
- Cálculo da multiplicação  $C = FFT(A) * FFT(B)$
- Cálculo da 2D\_FFT inversa de C

O cálculo da 2D-FFT de uma matriz  $A(M \times N)$  é realizado através dos passos:

- Cálculo de  $N$  x 1D\_FFT de comprimento  $M$
- Cálculo da transposta da matriz FFT(A)
- Cálculo de  $M$  x 1D\_FFT de comprimento  $N$

A implementação paralela baseada em paralelismo de dados usando HPF apresentou os resultados mostrados na Figura 1, para matrizes de dimensões (512x512) e (1024x1024). Neste gráfico,  $np$  é o número de processadores,  $t_{comp}$  o tempo de computação e  $t_{comm}$  o tempo de comunicação.

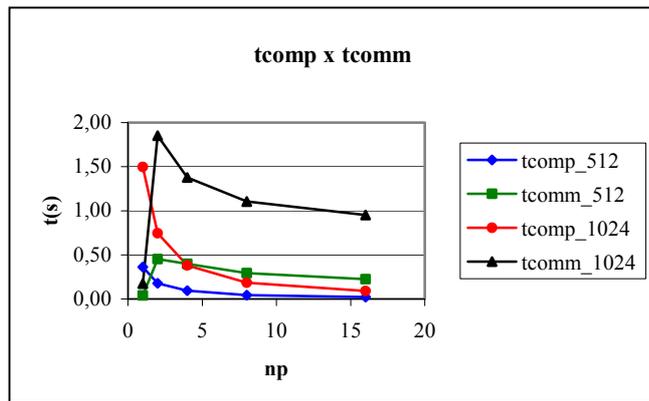


Figura 1 - Resultados parciais

Sendo as matrizes distribuídas em blocos de colunas entre os processadores, a comunicação gerada na paralelização é devida ao tempo de transposição de matrizes. Com o intuito de minimizar o tempo de comunicação, estão sendo pesquisadas técnicas alternativas para implementação deste algoritmo. Uma técnica em estudo consiste em utilizar diretivas HPF para distribuição dos dados e funções MPI para a transposição de matrizes, exemplificada na Figura 2.

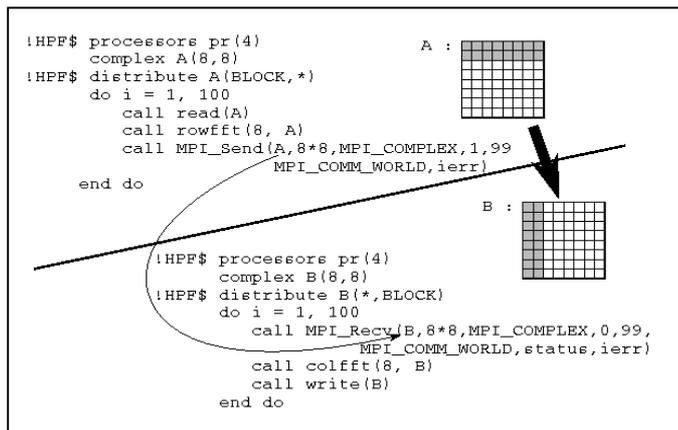


Figura 2 - Implementação HPF / MPI

### Referências

- [1] Wright, E.L. *Scanning and Mapping Strategies for CMB Experiments*. The Astrophysical Journal, astro-ph/9612006, 1996
- [2] Borrill, J., *The Challenge Of Data Analysis For Future CMB Observations*. The Astrophysical Journal, astro-ph/9903204, 1999
- [3] Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes in Fortran 90 - The Art of Parallel Scientific Computing*. Cambridge University Press, 1996
- [4] Koelbel, C.H.; Loveman, D.B.; Schreiber, R.S.; Steele Jr., G.L.; Zosel, M.E. *The High Performance Fortran Handbook*. The MIT Press, 1997
- [5] Snir, M.; Otto, S.; Huss-Lederman, S.; Walker, D.; Dongarra, J. *MPI - The Complete Reference, Volume 1 - The MPI Core*. The MIT Press, 2000
- [6] Foster, I.; Kohr Jr., D.R.; Krishnaiyer, R.; Chouldhary, A. *Double Standards: Bringing Task Parallelism to HPF Via the Message Passing Interface*. In: SC96 Technical Papers, 1996.